



TCB No. 004-2009

April 2009

Technical Bulletin

Genome Sequencer FLX System

Using Multiplex Identifier (MID) Adaptors for the GS FLX Titanium Chemistry - Basic MID Set

Summary

This bulletin describes the use of a set of ten Multiplex Identifiers (MID) with the GS FLX Titanium chemistry. This document includes information on ordering and preparing MID Adaptors as well as instructions for their use in library preparation. Libraries prepared with these Adaptors may be multiplexed in emulsion PCR (emPCR) in order to facilitate sample identification following sequencing on the Genome Sequencer FLX Instrument. This enables the most efficient sequencing for the user and provides the flexibility to combine gasket format with MIDs as needed to meet one's particular multiplexing needs.



Note: The instructions provided in this document negate the statement in the GS FLX Titanium General Library Preparation Method Manual to the effect that the GS FLX Titanium chemistry does not support the use of MIDs

Introduction

The information contained in this document is provided to enable users of the GS FLX Titanium sequencing chemistry to employ Multiplex Identifier (MID)-containing adaptors for General (e.g. Shotgun) library preparation. Please note that the *GS FLX Titanium General Library Preparation Method Manual* and the GS FLX Titanium General Library Preparation Kit can be used to prepare a library without reference to the information contained in this bulletin. If you are preparing standard, non-MID libraries, you do not need this document.

This document describes the preparation and use of up to ten Multiplex Identifier (MID) Adaptors. These may be used as a replacement for the Adaptors provided in the GS FLX Titanium General Library Preparation Kit. The MID Adaptors include a 10-nucleotide sequence tag on Adaptor A which is unique for each MID. When different libraries are prepared with different MIDs, they can be amplified by emPCR and sequenced together, in a multiplex fashion; the sequencing reads can be deconvoluted by the data analysis software after the sequencing Run, such that the reads from each of the pooled libraries are identified by their MID tag and correctly assigned.



Note: The two different chemistries of the Genome Sequencer System (GS FLX standard and GS FLX Titanium) are not compatible with each other. For example, the GS Multiplex Identifiers (MID) 1–6 Kit and the GS Multiplex Identifiers (MID) 7–12 Kit cannot be used with GS FLX Titanium series reagents. GS FLX standard series libraries cannot be amplified using GS FLX Titanium series emPCR kits, and libraries amplified with GS FLX Titanium series kits cannot be sequenced on the GS FLX standard series PicoTiterPlate devices, or vice-versa.

Figure 1 illustrates the structure of a completed library DNA fragment with and without an MID Adaptor:

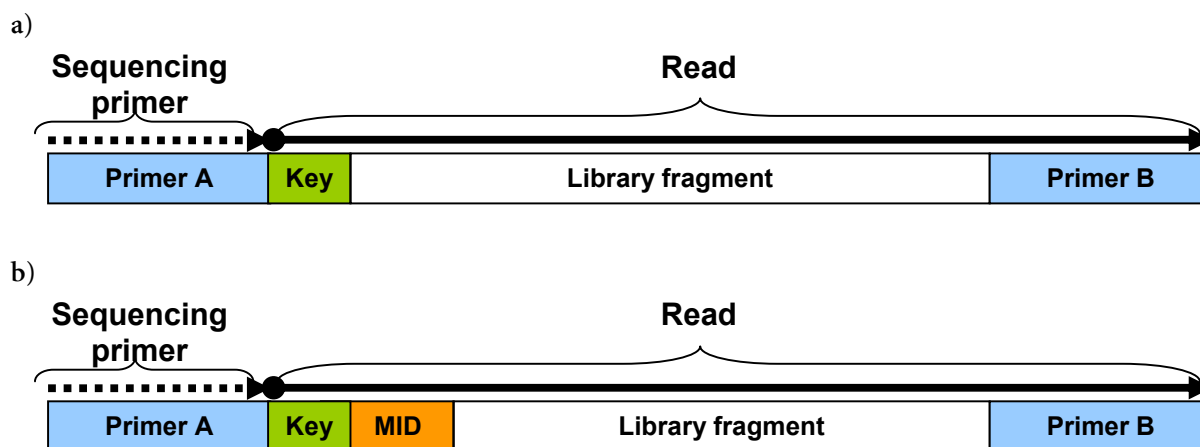


Figure 1. Schematic of final library constructs with regular Adaptors (a) and with MID Adaptors (b).

The MID sequence is added to 3' end of the Adaptor A sequence, immediately following the four-base Key. Because of this, the new Adaptor A must be synthesized with the addition of the desired MID sequence. Although only Adaptor A is new, both Adaptors are mixed in the Library kit, so all oligomers in Adaptor A and B must be synthesized and prepared as a new mixture for each MID.

Multiplex Identifier Adaptor Design

A robust set of ten decamer Multiplex Identifier (MID) sequences was designed to facilitate library multiplexing in 454 Sequencing. A length of ten bases is sufficient to ascertain that, for the large number of reads involved and the design parameters considered, the chances of mis-assigning reads is extremely low (see below). MID-containing adaptors can be incorporated during library preparation, yielding libraries where each fragment contains its specific library-identifying sequence. Different MID-labeled libraries can be subsequently mixed during emPCR and sequencing, and the resulting library reads can be reliably sorted after sequencing using SFF software tools as outlined in the *Genome Sequencer Data Analysis Software Manual*.

General (shotgun) libraries can be prepared using MID-containing adaptors in place of the Adaptors supplied in the GS FLX Titanium General Library Preparation Kit. As mentioned above, the standard GS Multiplex Identifiers (MID) kits CANNOT be used in the GS FLX Titanium chemistry. When used as a set of ten, the ten MID's described below allow extensive error correction capability as well as robustness to avoid read mis-assignment. Error is defined as a single insertion, deletion, or substitution.

The ten MID's were designed so that they:

- Use at most 5 flow cycles
- Are all of the same length, but short enough to allow accurate oligonucleotide synthesis
- Avoid monomer repeats
- Do not form a repeat with last base of the Key sequence (G)
- Are robust to noise (can correct 2 and detect 3 errors)
- Produce nonsense if two different MID's are read simultaneously off the same bead.

One critical element of the design is robustness against deletions, which are common during oligonucleotide synthesis. For example, two sequences like ACTGA and CTGAG may seem far apart when considering substitutions, but they can become one base apart or even identical with only a single deletion of the 5'A of ACTGA.

The set of ten 10-base MID sequences are shown in Table 1 below. These MID sequences are already included in the MIDConfig.parse file described in the *Genome Sequencer Data Analysis Software Manual*, section 13.4.15 (MID Configuration File). The ten-nucleotide identifier portion of these MID sequences are identical to those of the same name from the GS FLX Standard Multiplex Identifiers (MID) 1-6 and 7-12 Kits.

Table 1. 10-base (MIDs) Multiplex Identifier Sequences

MID-1	ACGAGTGCGT
MID-2	ACGCTCGACA
MID-3	AGACGCACTC
MID-4	AGCACTGTAG
MID-5	ATCAGACACG
MID-6	ATATCGCGAG
MID-7	CGTGTCTCTA
MID-8	CTCGCGTGTC
MID-10	TCTCTATGCG
MID-11	TGATACGTCT

Obtaining and Preparing MID Adaptors

Ordering Oligonucleotides



Note: The oligonucleotides required for adaptor annealing may be ordered from a variety of suppliers. Screenshots provided herein were obtained from Integrated DNA Technologies (IDT - <http://www.idtdna.com>) and are meant to serve as an example of the process one would be expected to follow.

1. For each different MID desired you must obtain the required oligonucleotides and prepare a new Adaptors mix. Each MID Adaptors mix is comprised of an Adaptor A and an Adaptor B. Each adaptor is comprised of two oligonucleotides that are annealed in an equimolar ratio as described in the next section. These adaptors are then combined to make an MID Adaptors mix.
2. The list of oligonucleotides required to make all ten MID Adaptors A plus the common Adaptor B are provided in the Appendix. For example, to make the MID1 Adaptor mix, oligonucleotides "Ti-MID-A", "Ti-MID-Aprime", "Ti-MID-B" and "Ti-MID-Bprime" must be obtained as shown in the Appendix.



3. Oligonucleotide synthesis specifications:
 - a. Each oligonucleotide should have phosphorothioate modifications in both the first four and last four bases of the oligomers.
 - b. The Adaptor B long oligonucleotide 'Ti-MID-B' must be synthesized with a 5-prime biotin-TEG moiety.
 - c. All oligonucleotides must be purified using HPLC.
4. Choose a synthesis scale that yields the appropriate mass of oligonucleotide for your needs. Consult your supplier of choice to understand the minimum guaranteed yield at a given synthesis scale.
 - a. For example, at a 100 nanomole scale with HPLC purification, IDT's guaranteed yield is at least 2 nanomoles for all "A" and "Aprime" oligonucleotides. This is sufficient material to follow the workflow specified in the next section whereby 2 nanomoles of each oligonucleotide are used in annealing. You will need a greater amount of the Adaptor B oligonucleotides as it is common to all MID Adaptors mixes (i.e. Adaptor A contains the MID and Adaptor B is common). At a 1 μ mole synthesis scale, IDT's guaranteed yield for Adaptor B oligonucleotides is greater than 35 nanomoles. This is sufficient to make one preparation each of all ten MID Adaptor mixes as outlined in the next section.
5. See Figure 2 on the next page for a graphical representation of these ordering specifications using the four oligonucleotides required to make the MID1 Adaptors mix. This figure represents a screenshot from the IDT Oligonucleotide ordering screen.
6. Oligonucleotides may be ordered lyophilized (dry) or in solution (generally TE, pH 8.0) at a fixed concentration (i.e. 100 μ M). Store these materials per the supplier's recommendations. Generally, either room temperature or -20°C storage is appropriate for lyophilized material, and -20°C is specified for oligonucleotides in solution.



Order Oligos

Change Form: Expand to this many items ☐ Duplex ☐ Paste

25 nmole DNA oligo = 15-60 bases 100 nmole DNA oligo = 10-90 bases
 250 nmole DNA oligo = 5-100 bases 1 µmole DNA oligo = 5-100 bases
 5 µmole DNA oligo = 5-60 bases 10 µmole DNA oligo = 5-60 bases
 25 nmole Ultramer™ DNA Oligo = 60-200 bases 4 nmole Ultramer™ DNA Oligo = 60-200 bases
 PAGE Ultramer™ DNA Oligo = 60-200 bases

Change Scale: 100 nmole DNA oligo

Change Purification: HPLC \$42.00

Scale: 100 nmole DNA oligo **Purification:** HPLC \$42.00

Sequence Name Ti-MID1-A **# Bases:** 40

5'-C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGACGAG
T*G*C*G*T-3'

Notes: GS FLX Titanium MID1 A oligo

Scale: 100 nmole DNA oligo **Purification:** HPLC \$42.00

Sequence Name Ti-MID1-Aprime **# Bases:** 22

5'-A*C*G*C*ACTCGTCTGAGTCG*G*A*G*A-3'

Notes: GS FLX Titanium MID1 A-prime oligo

Scale: 1 µmole DNA oligo **Purification:** HPLC \$105.00

Sequence Name Ti-MID-B **# Bases:** 30

5'-/5BioTEG/C*C*T*A*TCCCCTGTGTGCCTTGGCAGTC*T*C*A*G-3'

Notes: GS FLX Titanium B oligo
(common to all MID Adaptor mixes)

Scale: 1 µmole DNA oligo **Purification:** HPLC \$105.00

Sequence Name Ti-MID-Bprime **# Bases:** 12

5'-C*T*G*A*GACT*G*C*C*A-3'

Notes: GS FLX Titanium B-prime oligo
(common to all MID Adaptor mixes)

Figure 2. Screenshot of IDT oligonucleotide ordering screen. In this example, four oligonucleotides are shown, as required to make the MID-1 Adaptors mix. Note that HPLC purification is specified for all oligonucleotides. Also shown are the required phosphorothioate modified bases and 5'-Biotin-TEG moiety on the 'Ti-MID-B' oligonucleotide. Note the different synthesis scales as described above.

Annealing Adaptors



Note: Each MID Adaptors mix (mixture of Adaptors A and B for a given 10-base MID) must be prepared individually. The instructions below outline the process of making one such Adaptors mix (MID-1), in which Adaptors A and B are separately annealed, then mixed to form the resultant MID Adaptors mix. Examples of annealed Adaptors A and B are provided in the Appendix.



Shelf life: 1× TE and Oligo Buffer have a shelf life of 3 months; indicate the expiration date on the container, and do not use after the expiration date has passed.

1. Obtain or prepare 1X TE Buffer (10mM Tris, pH 8.0, 1 mM EDTA)

49.5 ml Molecular Biology Grade Water (many possible)	
<u>0.5 ml 100× TE, pH 8.0</u>	(e.g. Fisher Scientific BP1338-1)
50.0 ml total	

- a. Combine the above ingredients and mix thoroughly.
 - b. Store at −15 to −25°C (or temporarily at +2 to +8°C).
2. If oligonucleotides are lyophilized (dry), briefly centrifuge the tubes to ensure that any powdered material is concentrated in the bottom of the tubes.
 3. Dissolve or dilute each lyophilized oligonucleotide to a concentration of 100 μM with 1X TE. This is easily done by adding TE at ten times the number of nanomoles listed on the oligonucleotide spec sheet or based on your own OD₂₆₀ quantitation. For example, adding 20 μl of TE to 2 nanomoles of lyophilized oligonucleotide will yield a 100 μM stock solution. Mix by vortexing. Incubate at room temperature for five minutes before proceeding.
 4. (Optional) Verify the concentration of the oligonucleotide stock solution via OD₂₆₀ spectrophotometry or fluorometry. Guidance on this process may be found on the IDT website (link below) or as obtained from your oligonucleotide supplier of choice:

http://idtdna.com/Support/Technical/TechnicalBulletinPDF/Oligonucleotide_Yield_Resuspension_and_Storage.pdf

5. Prepare Oligo Buffer (1X TE and 50 mM NaCl) as follows:

9.9 ml 1X TE, pH 8.0	(recipe above)
<u>0.1ml 5 M NaCl</u>	(e.g. Invitrogen 24740-011)
10 ml total	

- a. Combine the above ingredients and mix thoroughly.
- b. Store at −15 to −25°C (or temporarily at +2 to +8°C).

6. Oligonucleotide matching: Mix the complementary oligonucleotides A1 with Aprime1 in Oligo Buffer following the instructions in Step 7 below. Then mix oligonucleotides B1 with Bprime1 in a second, separate tube.



Caution: Do not mix the oligonucleotides for Adaptor A and Adaptor B together prior to annealing! The two oligonucleotides for Adaptor A must be annealed in one tube and the two oligonucleotides for Adaptor B are annealed in a second, separate tube.

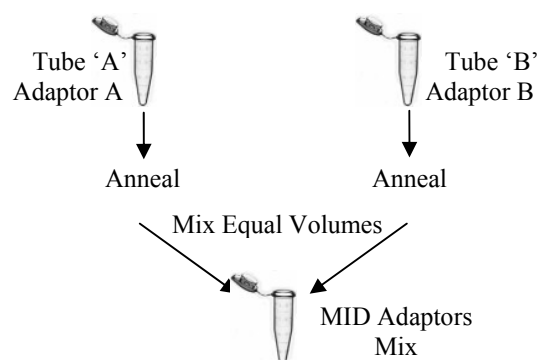
7. Add resuspended oligonucleotides and Oligo Buffer into 0.2 ml tubes to make the Adaptor A and Adaptor B annealing reactions for the MID mix being prepared. Label each tube before transferring any liquid to prevent errors. The schematic below demonstrates the creation of MID-1 Adaptors mix:

Adaptor A (Tube 'A')

60 µl Oligo Buffer
 20 µl Ti-MID1-A oligo (100 µM)
20 µl Ti-MID1-Aprime oligo (100 µM)
 100 µl total

Adaptor B (Tube 'B')

60 µl Oligo Buffer
 20 µl Ti-MID-B oligo (100 µM)
20 µl Ti-MID-Bprime oligo (100 µM)
 100 µl total



8. Mix each Tube 'A' and 'B' by vortexing, and spin each down briefly. Place the two 0.2 ml tubes in a thermocycler and run the following annealing program:
 - a. 95°C for 1 minute
 - b. Decrease temperature by 0.1°C / second to 15°C final temperature
 - c. 14°C forever
9. Combine an equal volume of each annealed Adaptor A and B (e.g. 100 µl of each using the above example) to prepare the MID Adaptors mix. Mix by vortexing. The final concentration of each Adaptor in this mix is 10 µM.
10. The MID Adaptors mix will be used as a substitute for the Adaptors supplied in the GS FLX Titanium General Library Preparation Kit, as described in section 3.5 (Adaptor Ligation) of the *GS FLX Titanium General Library Preparation Method Manual*. The adaptor ligation step calls for 5 µl of the final Adaptor mix per library, so enough MID Adaptors mix has been prepared to make up to 40 libraries following the above protocol.
11. Distribute 50 µl aliquots (each aliquot will be sufficient for up to 10 libraries) of the MID Adaptors mix into fresh tubes for storage. Label aliquots with the full (e.g. 'MID-1') name of the Adaptors mix.
12. Store the MID Adaptors mix aliquots and any remaining separate Adaptor stocks at -15 to -25°C. These materials may be stored under these conditions for up to one year.

Preparing and Sequencing Libraries with MID Adaptors

1. As specified above, the MID Adaptors mix of interest will be used as a substitute for the (non-MID)Adaptors in section 3.5 (Adaptor Ligation) of the *GS FLX Titanium General Library Preparation Method Manual*. As prescribed in the manual, 5 µl of the desired MID Adaptors mix should be substituted for the Adaptors supplied in the GS FLX Titanium General Library Preparation Kit.
2. Except for the substitution of the MID Adaptors mix for the supplied Adaptors, all steps prior and subsequent to adaptor ligation in the General library preparation should be carried out according to the *GS FLX Titanium General Library Preparation Method Manual*.
3. Following library synthesis with MID Adaptors, quality assessment and quantitation should be performed according to the procedures specified in Section 3.10 of the *GS FLX Titanium General Library Preparation Method Manual*.
4. Subsequently, libraries should be diluted and stored as prescribed, and the appropriate amount of library to be used in emPCR amplification must be determined by one of the methods specified in Section 3.10 of the *GS FLX Titanium General Library Preparation Method Manual*.
5. Libraries prepared with MID Adaptors may be amplified using either the *GS FLX Titanium LV emPCR Kit (Lib-L)* or the *GS FLX Titanium SV emPCR Kit (Lib-L)* products. As a reminder, libraries prepared with MID Adaptors are not compatible with GS FLX standard series reagents – only Titanium series.



Note: When preparing multiple MID libraries to be pooled and sequenced together, it is not always necessary to perform the functional quantitation procedure on each library separately. Instead, equimolar amounts of each library can be pooled and the titration carried out on the pool. Be aware, however, that the pooled libraries may perform unevenly, and that some may even fall outside the specifications of the system.

6. Enriched beads obtained following emPCR amplification of MID adapted libraries should be sequenced according to the *GS FLX Titanium Sequencing Method Manual*. No changes in bead loading or at any other steps in the protocol are required; handling is identical to that of non-MID General (shotgun) library beads.
7. Sequencing data should be analyzed according to the instructions supplied in the *Genome Sequencer Data Analysis Software Manual*. Directions are provided for the analysis of MID adapted library data using the GS De Novo Assembler (Section 5) and GS Reference Mapper (Section 6) applications, as well as SFF Tool Commands (Section 7).

Appendix

MID Oligonucleotide Sequences for Ordering

OLIGO NAME	OLIGO SEQUENCE (5-prime to 3-prime orientation)
Ti-MID1-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGACGAGT*G*C*G*T
Ti-MID2-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGACGCTC*G*A*C*A
Ti-MID3-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGAGACGC*A*C*T*C
Ti-MID4-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGAGCACT*G*T*A*G
Ti-MID5-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGATCAGA*C*A*C*G
Ti-MID6-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGATATCG*C*G*A*G
Ti-MID7-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGCGTGTCT*T*C*T*A
Ti-MID8-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGCTCGCG*T*G*T*C
Ti-MID10-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGTCTCTA*T*G*C*G
Ti-MID11-A	C*C*A*T*CTCATCCCTGCGTGTCTCCGACTCAGTGATAC*G*T*C*T
Ti-MID1-Aprime	A*C*G*C*ACTCGTCTGAGTCG*G*A*G*A
Ti-MID2-Aprime	T*G*T*C*GAGCGTCTGAGTCG*G*A*G*A
Ti-MID3-Aprime	G*A*G*T*GCGTCTCTGAGTCG*G*A*G*A
Ti-MID4-Aprime	C*T*A*C*AGTGCTCTGAGTCG*G*A*G*A
Ti-MID5-Aprime	C*G*T*G*TCTGATCTGAGTCG*G*A*G*A
Ti-MID6-Aprime	C*T*C*G*CGATATCTGAGTCG*G*A*G*A
Ti-MID7-Aprime	T*A*G*A*GACACGCTGAGTCG*G*A*G*A
Ti-MID8-Aprime	G*A*C*A*CGCGAGCTGAGTCG*G*A*G*A
Ti-MID10-Aprime	C*G*C*A*TAGAGACTGAGTCG*G*A*G*A
Ti-MID11-Aprime	A*G*A*C*GTATCACTGAGTCG*G*A*G*A
Ti-MID-B	/5BioTEG/C*C*T*A*TCCCCTGTGTGCCTTGGCAGTC*T*C*A*G
Ti-MID-Bprime	C*T*G*A*GACT*G*C*C*A

- Phosphorothioate bonds are indicated with an asterisk (*)
- A 5'-biotin-TEG moiety is indicated by '/5BioTEG/'



Purification: All oligonucleotides must be ordered with HPLC purification and with the modifications (i.e. phosphorothioate bonds and 5'-biotin-TEG) shown.



Examples of Annealed MID Adaptors

MID-1 Adaptor A:

5' -CCATCTCATCCCTGCGTGTCTCCGAC**TCAG****ACGAGTGC**GT-3' Ti-MID1-A
3' -AGAGGCTG**AGTC****TGCTCACGCA**-5' Ti-MID1-Aprime

MID Adaptor B (Common):

5' Biotin-TEG-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG-3' Ti-MID-B
3' - ACCGTCAGAGTC-5' Ti-MID-Bprime

- Phosphorothioate-modified bases are not shown in this figure for ease of sequence alignment
- Sequencing Key is indicated in **GREEN**. 10 bp MID sequence is shown in **YELLOW**.